DOCUMENT RESUME

ED 159 286 UD 018 585

AUTHOR Menges, Constantine Christopher
TITLE Knowledge and Action: The Use of Social Science Evaluations in Decisions on Equal Educational Opportunity, 1970-73.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE 78
NOTE 73p.

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS Compensatory Education Programs; Data; *Decision Making; *Educational Policy; Elementary Secondary Education; *Equal Education; *Federal Programs; Government Role; *Information Utilization; Integration Effects; Needs Assessment; Policy Formation; *Program Evaluation; School Integration; Social Science Research

ABSTRACT

This report analyzes why the Federal government has made little use of evaluation studies and data available on equal education and school integration. An overview is provided of Federal policy and action on equal education from 1970 to 1972. The utilization of evaluations in the areas of educational and social effects of desegregation, criteria/incentives for maximum feasible desegregation, and compensatory education and achievement are discussed. Also traced are political, bureaucratic, and intellectual factors which affected the extent to which evaluations shaped Federal policy. It is held that with more communication between legislative staff and researchers, valid social program evaluations can be chosen and cumulatively synthesized, in order to apply them to decision making and policy. (Author/KR)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

# KNOWLEDGE AND ACTION: THE USE OF SOCIAL SCIENCE EVALUATIONS IN DECISIONS ON EQUAL EDUCATIONAL OPPORTUNITY, 1970-73

ED159286

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

By: Dr. Constantine Christopher Menges
SENIOR FELLOW,
NATIONAL INSTITUTE OF EDUCATION

For: The National Institute of Education
Washington, D.C.

1978

UD018585

AUG 30 1978

TABLE OF CONTENTS


| | Page |
|---|---|
| INTRODUCTION | |
| A. FEDERAL POLICY AND ACTION ON EQUAL EDUCATIONAL OPPORTUNITY-AN OVERVIEW 1970-72 | 1 |
| B. THREE MAJOR ANALYTIC THEMES FOR DECISION | 10 |
| C. THE USES MADE OF EVALUATIONS | 14 |
| I SOCIAL AND EDUCATIONAL EFFECTS OF SCHOOL DESEGREGATION | 18 |
| II CRITERIA/INCENTIVES FOR MAXIMUM FEASIBLE DESEGREGATION | 26 |
| III COMPENSATORY EDUCATION AND ACHIEVEMENT | 37 |
| D. FACTORS AFFECTING THE USE OF EVALUATIONS | 51 |
| E. RECOMMENDATIONS FOR MEANS TO LINK EVALUATIONS AND DECISIONS MORE EFFECTIVELY | 57 |
| FOOTNOTES | 66 |
| BIBLIOGRAPHY | 68 |

Knowledge and action-the use of social science evaluations for national decisions on equal educational opportunity, 1970 to 1973

INTRODUCTION

Since the mid-1960's, the federal government has greatly expanded its commitment to the solution of major social problems. Major new programs were launched in health, education, housing, manpower and other social services. Federal expenditures for social programs increased in real and proportionate terms. In 1950 the federal government spent 23% of its budget on social purposes, this increased to 28% in 1970 and 56% or $170 billion by 1975.(1) This growth in federal effort was matched by heightened public attention to the problems of the least well off Americans and increased moral concern for their needs.

As these new federal programs were established, the Congress often included a requirement that the executive branch conduct an independent and systematic evaluation of the results actually attained. This evaluation requirement reflected the fact that the Congress had seen too many well intentioned laws accomplish little or go astray. It also reflected an understanding that these social programs were moving into difficult and complex domains of American life and would probably require thoughtful adjustment based on experience. The request for evaluation also expressed confidence that applied social science methods could produce objective and meaningful results. It also recognized that the usual reporting methods available to executive branch managers had not been widely successful. There is after all, a built in natural incentive for a federal agency to highlight the positive in reporting on the impact of its programs for many reasons including the fierce annual competition for budget and staff resources.

During the years of expanding social programs, the federal government has spent many hundreds of millions of dollars on social science evaluations. In 1968 only $20 million was spent

for evaluations throughout the federal government. But, since 1970 the Department of Health, Education and Welfare (HEW) alone has provided more than $60 million each year for this purpose. During the first five years of the 1970s, HEW contracted for 160 evaluation studies in education and 194 in health programs. This large investment has produced reports which vary greatly in method and quality. Some evaluations try to measure the impact of programs in a particular year, some are long term studies with carefully chosen representative samples. Some use all the most advanced social science techniques such as control groups and highly quantitative analysis of data, others are a type of "natural experiment" and many are organized reportage and observation.

One generalization applying to all these evaluation studies, however, is that far too little use has actually been made of them inside the federal government. The very sensible and indeed courageous commitment of a democratic government to provide for objective assessments of its major social programs has usually not produced the linkage between objective knowledge and action which had been intended. As early in the process as 1970, a review of federal evaluation efforts concluded that these had no impact on national policy.(2) The early optimism of the late 1960s, that evaluation could add a valuable new dimension to government policy making has been replaced by ever growing doubts among members of Congress, the executive branch and the public about whether this investment will ever produce results which have a tangible impact on what government does.

This analysis will suggest specific recommendations for actions which can help bridge the wide gap between evaluations and executive decision-making. These suggestions will derive from a careful examination of the way evaluations were used and ignored during a three per year period when HEW took major decisions on equal educational opportunity. This case study

approach will provide a realistic sense for the character of executive branch action in an important policy area so that the recommendations will be relevant to the real world rather than the neat and clean patterns of the organization charts.

Beginning with a speech by the President in 1970, the federal government began a major effort to provide help for public school districts undertaking desegregation. The first section of this report will provide an overview of the background and actions taken by the federal government during a three year period of important decisions, 1970-73. The second part will describe the three major policy issues which emerged from the initial analytic effort in 1970 and which remained as the leading questions throughout the administration.

Next, this report will trace the use made of social science evaluations in an effort to shape policy on each of these three principal questions. The political, bureaucratic, and intellectual factors which affected the extent to which the evaluations were used will be analyzed. A concluding section will offer some general observations on this review and practical suggestions for making better use of the evaluations which are currently available.

A final word of introduction about my own role in these events is required to help the reader place my judgements into perspective. I was invited to join the planning staff of the Secretary of HEW in March 1970 and had a major staff role in undertaking policy analysis on equal education opportunity issues during the time covered by this study. My responsibilities included the definition of many of the policy issues and alternatives which were acted upon by the Secretary and the President, the design and monitoring of evaluation studies and a good deal of analytic combat with the program leadership in the Office of Education and the White House. In looking back on a period of extremely intense and interesting professional work, I have tried to be as objective and thorough as human nature allows. Without question, however, other participants might interpret and analyze these events in quite different terms.

A. Federal action on equal educational opportunity, 1970 - 1973

In 1954 the Supreme Court ruled that the separation of children by race in schools was unconstitutional. Nevertheless, fourteen years later more than 80% of all southern black students were in highly segregated schools together with 55% of all other black students in the nation (see table 1). This slow pace of desegregation had persisted through the presidential administrations of both political parties. But by 1969, the combination of a new Supreme Court decision, legal action by civil rights groups and increased pressures on HEW for enforcement resulted in a rapidly increasing momentum. By 1973, dramatic progress had been made. Southern public schools were more integrated than those of the north and only 8% of all southern black students remained in highly segregated schools.

Table 1 Public school desegregation, 1954-73

(Proportion of black pupils in schools with more than 90% black enrollments)

| | South | US |
|---|---|---|
| 1954 | 97% | 65% |
| 1968 | 80% | 55% |
| 1973 | 8% | 40% |

Source: 1968, 1973, HEW, Office for Civil Rights annual survey;

During the entire time of the first Nixon administration, (1969-73) school desegregation and busing were the focus of heated national debate and emotional controversy. In 1969 the pace of court ordered urban busing and demands for more vigorous enforcement by HEW found the new administration divided and ambiguous. As a result the President decided to issue a major policy statement which would clarify his position and commit the federal government to provide federal help to make school desegregation work. In the March 1970 statement President Nixon said he wanted to "reaffirm my personal belief that the 1954 decision of the Supreme Court was right in both constitutional

and human terms." (3) The President proposed to spend more than $1.5 billion for three purposes: to help school districts undertaking federally required desegregation; to encourage and support voluntary integration programs; and, to improve the quality of education in newly integrated schools.

That presidential commitment launched HEW on a new course that ultimately led to the expenditure of more than $380 million for desegregation assistance during these three years and the allcoation of more than six million dollars for related evaluation studies. After three months of intensive staff work and a series of cabinet meetings on the issue, the President submitted the Emergency School Assistance Act (ESAA) to the Congress in May 1970. During that summer it became evident that there was little chance of the bill being passed in time to cope with the needs of desegregating school districts in the fall of 1970 and this resulted in a temporary program to accomplish some of the same purposes. Appropriately known as the Emergency School Assistance Program (ESAP) this effort was shaped entirely by HEW administrative decisions and was in effect for three school years (1970, 1971, 1972).

In late 1971 the looming national elections of 1972 and the public uproar and confusion about busing and desegregation made these issues political dynamite. With few exceptions, national leaders in both parties and across the spectrum of conservative and liberal politics took public positions opposed to busing for desegregation. At the same time millions of children in hundreds of school districts were required by federal courts to change schools for the purposes of desegregation. In this context, President Nixon considered proposing a constitutional amendment prohibiting mandatory busing for desegregation. Elliott Richardson, then Secretary of HEW opposed this and sought to define a middle ground on this issue. This led to another round of intense policy analysis effort in HEW and resulted in a new approach to equal educational

opportunity issues which was accepted by the President. In March 1972, the President sent the Equal Educational Opportunity Act to Congress as a complement to ESAA and parts of the existing compensatory education program. If passed, this would have been the single largest federal education program.

Table 2 Overview of equal educational opportunity decisions: 1970-72

| Date | ESAA Emergency School Assistance Act | ESAP Emergency School Assistance Program | EEOA Equal Educational Opportunity Act |
|---|---|---|---|
| March 1970 | Presidential speech | | |
| May 1970 | ESAA bill sent to Congress | | |
| July 1970 | | ESAP I planning begins for 1970-71 school yr. | |
| Spring 1971 | Preparation for implementation | Review of problems with ESAP-proposal of new approaches | |
| Summer 1971 | | ESAP II begins with new management system | |
| Jan-March 1972 | | Review of lessons from ESAP II | Analytic work for new bill |
| March 1972 | | | President submits EEOA to Congress |
| Summer 1972 | ESAA authorization is passed; | ESAP III begins | |
| December 1972 | ESAA appropriation passed by Congress | | |
| January 1973 | ESAA is implemented by HEW | | |

A bill proposing a moratorium on all further busing for desegregation until the Congress established defined standards of reasonableness was submitted together with EEOA. Ultimately the Congress failed to act on both of these new initiatives but did pass ESAA in June of 1972. Funds were authorized in

December of 1972 so that the program could be implemented in early 1973. In the meantime, the Emergency School Assistance Program used a variety of existing laws and funds from other federal sources to establish a small scale version of the ESAA program.

This emergency program became the focus for HEW efforts to assist school desegregation and served in effect to test many of the original assumptions, conclusions and judgements that had shaped the ESAA bill. From the start, there was a great deal of controversy within HEW about the character of the emergency program. A coalition of "southern strategy" political interests and the normal bureaucratic tendencies of funding agencies pressed for a general assistance style of administration which would have allocated the available funds among as many southern school districts as possible. The analytic viewpoint urged the use of these funds as an incentive for implementing meaningful desegregation and their concentration on districts which were making the greatest effort.

This emergency program functioned for three years and grew from $89 million in 1970 to $190 million in 1972. The annual evaluations of the program together with other evaluations which already were underway or which were designed specifically to answer questions about the desegregation process, were among the most important which this analysis will consider. As this brief history makes clear, HEW made two types of decisions on these issues during this three year period: decisions about national policy on school desegregation and decisions about how to implement a program to help schools with this task. There were two major rounds of policy decision: in the spring of 1970 when the initial approach was defined and during the early months of 1972 when a very tough battle was fought behind the scenes and a moderate approach was accepted by the President. Important decisions about implementing this program were taken in the summer of 1970; then again in the summer of 1971 when the lessons of the first year were used to improve the program in its

second year and, finally, during the summer of 1972 when there was an effort to prepare simultaneously either to implement the entire ESAA program or a third temporary program in September 1972 - depending on whether or not Congress appropriated the funds in time. On the whole, the evaluation results were much more often used to make decisions during the policy discussions than when implementation was at issue.

The Participants

There were essentially three clusters of participants in these decisions: the President and others working directly for his office; the Secretary of HEW and individuals working directly for him and the Commissioner of Education and his associates (see table 3). Decisions on policy which ultimately led to three Presidential bills being sent to the Congress tended to involve the White House far more than decisions on implementation. Policy decisions also tended to involve a far smaller group within HEW with the Secretary relying on his analytic staff to define the issues and alternatives and calling on the Commissioner of Education for reactions concerning feasibility. Implementation decisions involved a far wider grouping of people and the program managers in the Office of Education, quite naturally, had a strong voice in all aspects of this question.

Table 3 Executive branch - Participants

| The White House/Executive Office | |
|---|---|
| President Nixon | 1970-72 |
| Mr. John Erlichman | " |
| Mr. Edward Morgan | " |
| Mr. Paul O'Neil | 1972 |
| Mr. Patrick Moynihan | 1970 |
| Cabinet committee | 1970 |
| Mr. Mitchell | |
| Mr. Finch | |
| Mr. Agnew | |

| Department of H.E.W. | |
|---|---|
| Office of the Secretary | |
| Sec. Robert Finch | 1970 |
| Sec. E.L. Richardson | 1971-2 |
| ASPE-Mr. Lewis Butler | 1970-71 |
| Mr. Michael O'Keefe | 1970 |
| Mr. Michale Timpane | 1972 |
| Mr. C. C. Menges | 1970-72 |
| ASL Mr. Chris Cross | 1972 |
| GC Mr. Will Hastings | 1972 |
| ASM Mr. Backe | 1972 |
| Office of Education | |
| Comm. Terrence Bell | 1970 |
| Comm. Sidney P. Marland | 1971-72 |
| Assoc. C.Herman Goldberg | 1971-72 |
| Assoc. C. John Evans | 1971-72 |
| Prog.Mgr. William Rock | " |

This listing of participants indicates the extent of change in roles within this three year period. Only the White House and one member of the Secretary's planning staff in HEW participated throughout the entire time covered by this analysis. During the initial policy decisions in the spring of 1970 the majority of the work was done by the planning staff of Secretary Robert Finch. The 1972 policy decisions were taken in an environment of great secrecy due to the difficult battle being waged against the highly organized forces in the White House. This involved a very small group around H.E.W. Secretary Elliott Richardson which dealt with all the issues and consulted on specific items with a wider group which included the Office of Education.

All the participants viewed the decisions to be taken as important. They acted with widely different perceptions and policy purposes. The political explosiveness of school desegregation made these efforts to shape and define national policy and program incentives the object of careful and indepth attention by the President and his political advisors. Since tens of millions of dollars were made available by the ESAP program and promised by ESAA, the educational constituency demonstrated great interest in shaping the actual funding in ways most congenial with the priorities of school districts. For elementary and secondary education, this desegregation assistance program became the single largest source of new federal funds throughout the Republican administrations. Therefore, at all levels, the participants in these decisions believed that the outcome mattered very much.

B. Major Analytic Issues for Decision

The speech given by President Nixon on March 24, 1970 provided a comprehensive statement of the administration perspective on school desegregation. It affirmed support for the constitutionality of the historic 1954 Supreme Court decision, provided an overview of progress to date, surveyed the existing state of the law and then discussed some of the complexities posed by viewing this issue in the broader context of the education goals and social demography of public schools. Among the policies described under principles of enforcement were the following statements by the President:

- "deliberate racial segregation of pupils by official action is unlawful wherever it exists..it must be eliminated 'root and branch'--and it must be eliminated at once"
- "segregation of teachers must be eliminated"
- "school administrators throughout the nation..must move immediately..to assure that schools within individual districts do not discriminate with respect to...the quality of education delivered" (3)

Four categories of need were defined as the basis for the $1.5 billion program: desegregating districts; racial impacted schools where de facto racial isolation persisted; districts which "have the farthest to go to catch up educationally with the rest of the nation"; and, "innovative techniques for providing educationally sound inter-racial experiences for children in racially isolated schools." (3)

Within six weeks after this speech, HEW had formulated a series of alternative programs, these had been reviewed in depth by a special committee of the Cabinet and final decisions were made by the President which resulted in his submission of the Emergency School Aid Act (ESAA) to Congress in May 1970. It was during this time of policy analysis and development that the three principal issues which shaped decisions throughout the next years were crystalized.

In the spring of 1970, 664 southern school districts were either under federal court order or negotiating with HEW to begin desegregation in the fall of 1970. Another 634 southern school districts would be in their second year of desegregation that coming fall. In effect nearly all of the school districts in the south, were in the midst of a transition which would mark the end of their long defiance of the Supreme Court decision. This meant that more than fifteen million pupils, among the total public school enrollment of 46 million, were in districts facing federally mandated integration. For these school districts and children, the analytic issue was clear: what do the schools know about how to make desegregation effective in both educational and social terms? The President promised that "we shall launch a concerted, sustained and honest effort to assemble and evaluate the lessons of experience: to determine what methods of school desegregation have worked..and also what has not worked."(3) This became the first theme of analysis and evaluation efforts.

The new program was also intended to assist children where "de facto racial isolation persisted" and this led to a careful look at the racial demography of nation-wide public school enrollment. The Civil Rights Act of 1964 had provided legal authority for the Office for Civil Rights in HEW to collect annual data on school-by-school minority enrollment from a representative sample of school districts receiving federal aid which included nearly all minority students. In 1970 that data was used in an analytic way for the first time and revealed patterns that were quite surprising.

Total minority enrollment came to roughly 8.7 million pupils throughout the more than 18,000 school districts. But eight out of ten minority pupils were in 837 of these districts and nearly half were in the nation's 100 largest school districts. Segregation in the rest of the nation was nearly as pronounced as that in the south. In the fall of 1969 only 32% of minority

pupils were in majority white schools in the rest of the nation compared with 29% in the south. More than six out of ten black children were enrolled in virtually all black schools in that school year.

Further analysis of the data revealed that outside the south racial isolation was an especially severe problem in urban school districts. In Chicago 85% of all black children were in all black schools, in Los Angeles 78% were isolated in this way. The causes for this fact included housing patterns, spatial factors in big school districts and the changing racial demography of many large cities. By the 1969-70 school year the HEW analysis revealed that more than 28% of all minority children were in 19 urban school districts where minority children constituted more than half of the total enrollment. Examples included the District of Columbia-93% of all students were minority, Chicago-65%, New York-55%, and, Philadelphia, 59%. And the demographic patterns in these and many other large urban areas were increasing this trend. In fact, the trends suggested that most of the minority children in the fifty largest school districts would soon be in the majority in total enrollment. And, since these constituted 40% of all minority children, this called for immediate action. (see table 4 below)

Table 4 1969-70 Minority Enrollment in the 50 Largest School Districts

| % minority in public schools | number of districts | total enrollment (million) | Minority total (million) | % of all minority enrollment |
|---|---|---|---|---|
| More than 50% | 19 | 4 | 2.6 | 27% |
| Less than 50% | 31 | 3.3 | 1.1 | 12% |
| Totals | 50 | 7.7 | 3.7 | 39% |

Taking into account smaller school districts where minority pupils also constituted a majority led to the conclusion that for millions of minority children even a policy of immediate "proporational balance" in each school would leave them in racially isolated learning environments.

These facts, together with the distinction between illegal segregation resulting from official actions of local government and isolation resulting from housing patterns, led to the next two analytic themes. The federal courts were requiring desegregation in a number of cities which resulted in the additional expenditure of millions of dollars and the transportion of thousand of pupils to schools outside their neighborhoods. In Los Angeles, a state court, had ordered the school district to spend more than $100 million additional per year for busing which would have required more than 120,000 children to change schools. The political struggles sparked by desegregation were already leading to polarization between those who opposed any "forced busing" and those who viewed any effort to weigh the advantages and disadvantages of different levels of busing and integration as a pretext for continuing and condoning racial isolation. HEW moved into the difficult ground between these two absolute positions with the second theme for analysis: the search for criteria to measure and incentives to encourage the maximum feasible effort in desegregation, with a special emphasis on the complex situation of the large cities.

Enrollment patterns made it clear that any reasonable definitions of "maximum feasible effort" would still leave many minority children in racially isolated schools. Therefore there would have to be a search for efforts to improve the quality of education available to those children. And, there was a determination that newly integrating schools make special efforts to help children who might have been educationally disadvantaged. The Supreme

Court had found in 1954 that de jure segregated education nearly always meant an inferior education for minority children. For that reason, it was logical to assume that as children came together for the first time in the newly desegregating schools, there would be special need for compensatory help for some of those who had been in the schools of lower quality. These two purposes defined a third analytic theme: the educational effectiveness of compensatory education and means to make this approach to more equal opportunity work better.

Many other issues were of course considered in the spring of 1970 as the HEW initiative was being formulated. Whether or not to provide funds for voluntary as well as federally required desegregation - yes, was the answer. The regional distribution of the federal effort - whether there should be a special focus on the south or whether the program criteria should recognize the national scope of the issue. There were debates about funding methods--formulas or project grants, the role of community groups as means to provide a check on school boards which might not be fully committed to desegregation. And there was a successful effort to use this new program as an incentive for civil rights compliance by requiring this as a prior criterion for funding eligibility. But for all the many issues which were taken up in different substantive and procedural forms, these were the three analytic themes, which shaped the basic decisions of the government:

I. means to enhance the positive educational and social consequences of desegregation;

II. criteria and incentives for maximum feasible desegregation under a variety of school enrollment patterns;

III. the educational effectiveness of compensatory education and means to obtain more positive results

These issues emerged from the policy context, the facts of school social demography and the need to decide among different priorities in allocating the $1.5 billion federal effort.

C. The Use Made of Evaluations

The intense policy analysis effort which went into the search for a constructive program of desegregation assistance used whatever information and data was at hand, including evaluations. After the President had made his decisions and sent a bill to the Congress, the HEW analysts had the opportunity to reflect on the need for much better information about each of the three principal issues. The "quick and dirty" analytic work had shown the thinness and ambiguity of the social research which was available up to then.

Since it was evident that all three of these were large and enduring issues, I recommended that a portion of HEW evaluation resources be committed to finding better answers for the future. Several quite different factors worked to make it possible to undertake the evaluation studies listed below with the relevant analytic theme (see table 4):

Table 4 Overview-evaluations relevant to three analytic issues

| Analytic issue | Evaluation studies |
|---|---|
| I Effects of desegregation | a-HEW staff synthesis of evaluations,1970<br>b-Equal educational opportunity report 1966<br>c-ESAP I evaluation, 1971<br>d-Retrospective longitudinal study of achievement and desegregation,(1971)<br>e-Berkeley, Riverside Calif. evaluations of desegregation effects,1971,1972<br>f-ESAP II evaluation (1972)<br>g-HEW synthesis of evaluations,1972 |
| II Criteria/incentives for maximum feasible desegregation | a-Desegregation at different efficient levels of busing in 44 cities, 1971, 1972<br>b-Analysis approach to planning desegregation in a large city, 1972<br>c-Methods of urban desegregation planning without the use of computers, 1972 |

| | |
|---|---|
| III Effectiveness of compensatory education | a-Equal educational opportunity (Coleman report) 1966<br>b-HEW staff synthesis of evaluations, 1970<br>c-Reanalysis of title I achievement data, 1972<br>d-ESAP I evaluation, 1971<br>e-National Evaluation of Title I (1970-71), 1972<br>f-HEW synthesis of evaluation results, 1972<br>g-Longitudinal study of compensatory reading programs-designed and funded in 1971 |

HEW evaluation funds were allocated on the basis of a very conscious and systematic planning process. This resulted in the production of an annual evaluation plan which specified the priorities and evaluation budget for different programs or issues. The planning staff in the Office of the Secretary had overall review responsibility and by custom received about a quarter of the $14 million available for education evaluation at that time.

By May of 1970 the evaluation plan would normally have been in final form. However, the commitment of the President and the participation of many senior HEW officials including Secretary Finch and Undersecretary Veneman in the effort to frame and then explain the ESAA proposal to Congress provided support for an effort to find far better answers. Further, the ESAA bill required HEW to determine when a large urban district with a large minority enrollment had undertaken the "maximum feasible" desegregation and therefore was eligible for funding intended to benefit "racially isolated" students. The limitations of the first analytic look at the issues showed that HEW needed some help to be able to do this with any degree of objectivity. The concern for better information on factors which enhanced the success of school desegregation, had both an immediate and a medium range (1-3 years) time urgency. If the Congress passed the ESAA bill in the early summer of 1970, HEW

would have to implement it. And since the program was expected to continue for four years, there was an incentive for taking a deeper look at the issue and, above all, learning from the experience of the hundreds of school districts that would be desegregating for the first time in the fall of 1970. The same held for compensatory education programs.

This explains the three types of evaluation studies listed for these two issues. First, a synthesis of evaluation evidence available at the time fo decision-taking into account all available sources. Second, the effort to learn from recent educational experience in a more systematic way. Examples were the retrospective longitudinal study of desegregation and achievement and the reanalysis of title I achievement data which already existed. Third, the more typical evaluations of a new or ongoing program-for example the annual evaluations of ESAP I and ESAP II.

From a practical point of view it was very significant that the consensus about the need for evaluation studies which existed from the White House to the Office of Education, resulted in the HEW decision to allocate one percent of the emergency school aid program funds for this purpose. During these three years, this meant that nearly four million dollars was newly available for evaluation. As a result these desegregation related studies complemented rather than competed with other priorities. There was also an excellent spirit of cooperation and mutual support between the planning and evaluation staff in the Office of the Secretary and the Office of Education. The 1970 experience had convinced both that these three issues needed to be explored in greater depth and would be the focus of policy debate as the desegregation assistance program got underway.

I. Educational and Social Effectiveness of Desegregation

In April 1970, there were esentially two popular views about the consequences of school desegregation: it would lower educational standards for white children and accomplish nothing for minority children; or, it would always have positive effects for both groups. The search for evidence on this issue yielded a much more complex conclusion which pleased neither the opponents nor proponents of school desegregation.

A search of the education and evaluation literature produced surprisingly little evidence on the question of educational achievement and desegregation. The painstaking work of Meyer Weinberg in providing an overview of the available studies in this area (4) and the survey published in the Harvard Education Review (5) made it possible to come up to date with the existing studies. Twenty six educational studies and evaluations were judged to be reliable and used achievement test data. The conclusion was that in 12 cases school desegregation had positive effects and in 14 there was no change or negative achievement results. An effort to find school districts which had tried to evaluate the results of desegregation turned up a few examples-Berkeley, Riverside and Pasadena all in California. Their results were also mixed, tentative and only partially analyzed as of that time. The sum of this evidence placed HEW squarely between the extremes of popular opinion and suggested the need for further thought about the conditions under which integration might produce positive results.

In 1966 the Office of Education had published a report entitled Equal educational opportunity.(6)It was then and still remains probably the single best known study of education since the expansion of the federal role in the mid-1960s. Two main conclusions of the report were important in considering this issue: first,the large correlation between achievement and the socio-economic composition of the classroom; second, the fact that minority children performed better when schools were desegrated as long as more than half of their class mates were white.

The report viewed "minority" and "poor" as well as "white" and "middle income" as synonymous terms. Therefore, the conclusion about desegregation reinforced the primary conclusion of the report. (7)

Several immediate policy conclusions were drawn from this evidence. First, the clear view that desegregation was not only morally right and legally required in many circumstances, but that it could have positive results for all children. Second, that these positive results were not guaranteed but would have to be worked for by school officials. Third, there appeared to be no educational justification for desegregation that moved minority children from schools that were 95% minority in enrollment to those that were 50%, 60% or 70% minority in enrollment if these were mostly low income children.

These conclusions in turn led to the decision to emphasize the educational needs of desegregating school districts. They also were significant in the successful effort to persuade the President to offer support for desegregation that was entirely voluntary. Here the evidence that it could have positive results was essential. Finally, the pessimism about the educational benefits of moving children from one predominantly poor school to another in the large urban districts where minority children were the largest group led to the inclusion of a provision for educational assistance to school districts with more than 10,000 minority children or more than 50% minority enrollments.

Aside from those important consequences for the 1970 ESAA bill, the need for better evidence led to four new evaluation efforts (see Table 4, Items 1 -c,d,e) in 1970-71:

- a small contract for the analysis of the Berkeley, California longitudinal achievement data - before and during that district's voluntary desegregation program;

- a retrospective longitudinal analysis of changes in racial isolation and achievement patterns by ethnic group in seven to ten large cities;
- evaluation of educational outcome aspects of school desegregation in the South during the 1970-71 school year.
- Research design for a prospective longitudinal study of achievement and desegregation.

The first two of these projects were intended to provide information from existing desegregation situations. The results were to be available in the spring of 1971 when the first year of the "temporary" federal program would come under review. The fourth project, a research design, had a much longer time horizon. The difficulties in making any firm statements about the relationship between changes in racial enrollment patterns and achievement, prompted the Office of Education to consider a large scale, multi-million dollar effort to examine this issue provided a sound methodology could be developed. This contract for a careful methodologial design was to determine whether this could be done.

The ESAP I program evaluation was carefully designed to accomplish two purposes: to answer the policy questions raised by the 1970 analysis and provide information on the educational effectiveness of different techniques which school systems might use to make desegregation work. This evaluation was the first large scale systematic study ever undertaken of the school desegregation process. It involved a stratified random sample of all the school districts funded under the first year of ESAP. It included 252 districts in 14 southern states enrolling 51% of all the minority pupils in ESAP funded districts and 28% of all the minority pupils in those states. Using random sampling techniques in all phases, more than 9,000 project directors, principals, teachers and students of all ethnic groups were interviewed on a systematic and uniform basis by a multi-racial study team.

A large amount of desegregation actually occurred in the ESAP funded districts during the 1970-71 school year. One indicator of this is that the proportion of minority children in totally segregated schools declined from 98% in the 1969-70 school year to 30% by the fall of 1971. (8)

For various reasons it was not possible to obtain pupil achievement data for this evaluation. (9) Therefore, very careful planning was done to create the basis for a set of reliable measures of changes in relationships among faculty, administrators and pupils of different ethnic groups, as well as measures of changes in behavior related to the academic achievement of pupils. Perhaps the most significant finding of the evaluation was that the majority of the respondents felt that the racial climate had changed for the better or had not changed at all during the 1970-71 school year. Very few reported changes for the worse. Specific results are presented below in Table 5.

Table 5 Overview of the ESAP I Evaluation Results

| Area of Possible Change | Percent of teachers perceiving Situation as: Better | No change | Worse | Total |
|---|---|---|---|---|
| | % | % | % | |
| 1. The number of interracial friendships | 63 | 36 | 1 | 100% |
| 2. How well students of different races work together in class | 51 | 47 | 2 | 100% |
| 3. How well teachers of different races relate to each other | 34 | 63 | 3 | 100% |
| 4. The integration of student racial groupings on the campus and cafeteria | 32 | 66 | 2 | 100% |
| 5. The attendance of black students | 19 | 75 | 6 | 100% |

A preliminary report on these findings was available in the summer of 1971. At that time, HEW was assessing the results of the small ESAP I program as a basis for its planning for the implementation of the $1.5 billion ESAA which it was hoped would be passed by the Congress at any moment. The desegregation process evaluation was useful in providing a perspective on the rather widespread success of school desegregation and in emphasizing the fact that well-administered Federal assistance could have positive results.

However, the effort of the planning staff to argue that the program effectiveness conclusions of the evaluation should in some way be reflected in the funding priorities for ESAP II or ESAA went unheeded. The program manager contested the validity of the evaluation - it was a preliminary report and lacked achievement data. He refused to use the results directly. Yet, the evaluation report was among the factors spurring the program manager and his staff to attempt improvements in procedures used to assess the educational quality of desegregation assistance proposals submitted to HEW for funding.

By August of 1971 it had become apparent that ESAA would not be enacted by Congress in time for that school year. Once again Congress appropriated $75 million for a program of interim assistance which became known as ESAP II. In its second year, HEW made considerable managerial improvements in this program. In addition, the design for the evaluation of the second program incorporated many significant improvements in technique and also included the collection of achievement data. The basic focus of inquiry remained the same but the quality of information collected was better and, consequently, a more thorough evaluation was expected.

The 1972 Analysis

In early 1972, continued political attention was claimed by desegregation, especially in cities where the courts ordered extensive busing. This prompted the White House to begin planning for a constitutional amendment against all busing for integration and other measures to restrict busing. The Secretary of HEW, Elliott Richardson, decided to oppose the idea of a constitutional amendment and instead offer a constructive alternative.

In February 1972 an HEW staff synthesis was prepared summarizing what was known about the social and educational effects of desegregation. (10) The answer was still a mixed one, but this time the added evaluation material provided a far greater sense of confidence. The ESAP I evaluation made a positive case for desegregation, as did the Berkeley, California and Riverside California evaluations. This time, 22 of the 38 usable studies and evaluations available showed positive results.

A most unfortunate omission, was the retrospective longitudinal study which would have been especially valuable because it was intended to gather data from larger districts which might have been more representative of the problem areas in the nation. A series of problems encountered by the contractor led to this evaluation being terminated by the government(11).

The sum of this evidence was used to argue strongly that the administration should reinforce its commitment to desegregation assistance - both voluntary and required. It was used to counter the rising tide of congressional sentiment against further federal help for desegregation. Had this evaluation evidence not been available, it would have been extremely difficult to make a persuasive case concerning the overall positive educational results because the attention of the national media and politicians was focused on places where

desegregation was encountering great difficulty. Further, the short range political tide in both parties was running strongly against any encouragement of voluntary desegregation.

At the same time the mixed nature of the results, clearly suggested the need for special efforts in certain situations. This reinforced the policy conclusion reached in 1970, that extra educational help was necessary. Perhaps the best expression of the policy conclusions reached by the administration are the words of Elliott Richardson in testifying to the Senate on the Equal Educational Opportunity Act of 1972:

> "Sometimes desegregation improves the achievement of low income children, sometimes it does not....Any school district which genuinely seeks to solve its problems of educational deprivation will..achieve that part of the solution which desegregation can achieve by desegregating and achieve that part which will persist, because of poverty and environment. through good compensatory programs."(12)

In June 1972, Congress passed the Emergency School Assistance Act which had been submitted two years before and deferred further action on the new 1972 initiative. This meant that HEW now would have two years of experience with the smaller temporary program to use in making final preparations to implement the new law. The evaluation of the 1971-72 school year program had been carefully designed to provide even better knowlege about what helps make desegregation work. And, due to good leadership by the Office of Education planning and evaluation staff, the entire effort had been implemented on a time schedule which called for a preliminary analysis of the data to be completed by July 1972. This time the analysis staff intended to make certain that the program manager could not simply wave off the aspects of the evaluation results that were not in accord with his views. The final implementation decisions which were the subject of considerable dispute among various parts of HEW were to be made by the Secretary. I arranged for the timing of the final decision memorandum and meeting to be delayed until one week after the results of the new

evaluation were to arrive so that these could be used. In spite of protests and complaints about this delay-even though the Congress had not yet voted an appropriation for the new law! - this was done.

Unfortunately, the evaluation which would have had achievement test data from 24,000 students was late. The quest for better answers, the availability of a large evaluation budget and the spirit of controversy all combined to make the ESAP II evaluation a larger effort than that of the first year. A few days before the report was due, the contractor notified HEW that a computer problem involving lost or mixed up data tapes would delay the results for several months. The effect was, in my opinion, all the worse because there had been repeated promises that the results would be available on schedule. A great deal of political leverage had been used by the analytic staff to delay the final implementation decisions until the ESAP II evaluation was available. The immediate consequence was a loss of any major impact on the third year of the desegregation assistance program. But, perhaps just as important, the entire evaluation enterprise had been discredited in the eyes of HEW and Office of Education leaders who already were skeptical about any direct use of evaluation results anyway. Although it is beyond the scope of this study, my guess would be that the even more expensive and elaborate evaluation of ESAP III was even less used for subsequent management changes in the later and more expensive years of the program.

II. Criteria/incentives for Maximum Feasible Urban Desegregation

Until 1969 most desegregation had occurred in small school districts with enrollments less than a few thousand. As the federal courts became more active in urban cases, the complexity of the issue grew because tens and even hundreds of thousands of students were involved. If a court found illegal segregation in a large urban district, it would often prescribe some type of "proportional racial balance" as a solution. That is, if a district were 40% minority than students should be moved to new schools so that enrollments in all schools matched the racial composition of the district. But there was no constitutional or legal basis for this as a proper standard for remedy and the federal courts were in some disagreement among themselves on the issue. President Nixon contended in his March 1970 speech that this was not a criterion supported by anything in federal law up to that time. This still left the question-what are reasonable criteria for urban desegregation?

The 1970 school demographic analysis had shown that many minority children were in urban school districts where they were already in the majority. At the same time, when the surrounding metropolitan area was taken into account this pattern was reversed in all of the fifty largest school districts. This formed the basis for the special voluntary metropolitan desegregation incentives included in the 1970 ESAA bill. That bill also included a provision for education aid to racially isolated urban schools provided the district had made a reasonable effort to desegregate. This again required HEW to find some way of establishing criteria for making this judgment.

The issue was partly normative and partly empirical. But for anyone who held a normative position between those who would have ignored any but de jure segregation and those who

would have used proportional racial balance as a remedy in all cases, the definition of criteria would depend in part on costs and consequences. In the spring of 1970 it proved impossible to use existing information to come up with any useful criteria. Housing patterns, the spatial distribution of schools, local transportation networks together with varying proportions of minority enrollment all combined to create an incredibly difficult problem of assessment. For that reason, HEW designed an analysis approach to evaluate the extent of possible desegregation using different amounts of busing in urban areas.

One contractor analyzed desegregation possibilities in 44 cities and their metropolitan areas while another contractor dealt with the Los Angeles school district and its metropolitan area. Using data from a variety of existing sources, the following was done for each city: children enrolled in public schools were located by age and race; all public schools were located and their capacities and current grades recorded; a raltively detailed and authoritative transportation network including both distance and travel times was obtained and transposed into computer readable form. It was then possible to choose specific cases and parameters and simulate the assignment of children to schools in order to minimize transportation and costs while maximizing the extent of desegregation attained. This analysis used 35 minutes transportation time each way as the maximum, though this time parameter could be increased or decreased, and provided results for the following cases (as indicated by the X mark).

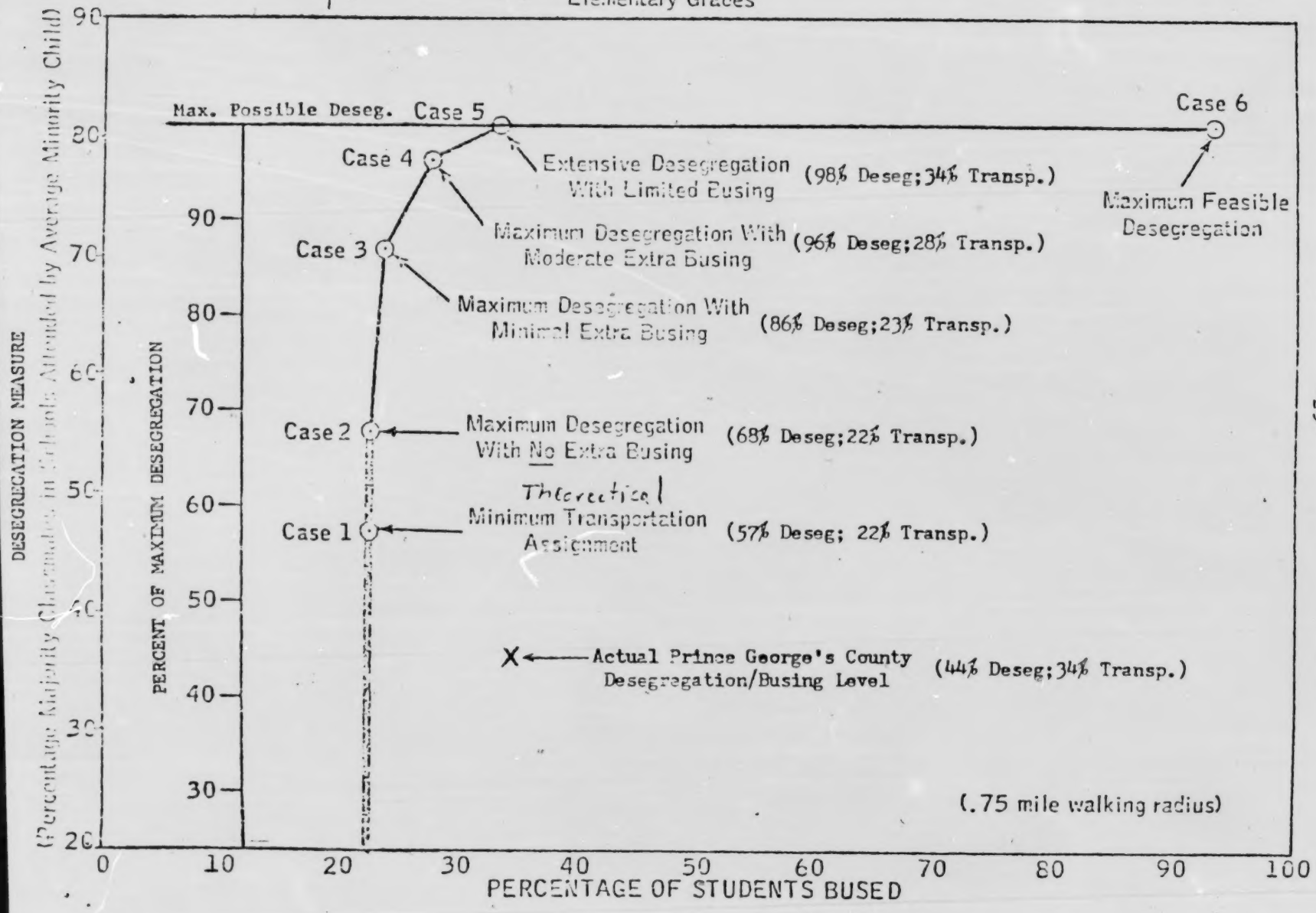
Table 6 Outline of Cases - Extent of Desegregation possible in

| Cases | the urban school district | the urban district and the surrounding metropolitan area |
|---|---|---|
| 1. Minimum transportation assignment/all children near enough walk to neighborhood schools. | X | X |
| 2. Minimum transportation assignment with an effort to desegregate. | X | |
| 3. Desegregation with minimal additional transportation (above the case #1 level). | X | |
| 4. Desegregation with moderate additional transportation. | X | |
| 5. Maximum feasible desegregation. | X | X |

The following illustration (table 7) provides an accurate portrayal of the six cases and results for one large school district. Note the very rapid improvement in desegregation outcomes possible with less than current levels of busing.

DESEGREGATION ANALYSIS RESULTS -- "A LARGE" COUNTY

Elementary Grades




-29-

TABLE 7

The preliminary results from the analysis showed that existing residential isolation of urban minority groups was not as serious a barrier to school desegregation as generally assumed. In many urban school districts significant reduction in racial isolation could be accomplished without busing any students who could otherwise walk to their neighborhood schools. And even in the largest cities analyzed, substantial school desegregation was possible without exceeding reasonable limits on travel time and with rather modest numbers of students transported. In many cities the analysis indicated that there could even be significant desegregation using current or even reduced amounts of busing. In other words, in these districts, by reassigning pupils to their neighborhood schools on a more efficient basis, there could be a simultaneous increase in desegregation and a decrease in busing.

For the cities and the metropolitan area the evidence was not as conclusive. It did reveal that in some of the major cities, desegregation of the entire metropolitan area could be accomplished while requiring the transportation of a smaller proportion of all the children than if the city alone were desegregating. In some other cities studied, metropolitan desegregation appeared to require transportation of a greater proportion of the children. In still others, the evidence was incomplete.

Due to the esoteric nature of this analysis, the techniques and conclusions have been presented at some length in order to make their use intelligible. In early 1972 the powerful political opposition to busing for desegregation could only have been countered by a coherent set of arguments. Fortunately the contractor for the 44 city analysis had completed the preliminary report on schedule-in December 1971. This enabled me to meet with Secretary Richardson early in January 1972 and brief him on the analysis, its methods, limitations and

the policy implications of the conclusions.

These were, first, that the use of existing or modest increases in busing would often permit substantial progress toward integration. Second, that the attempt to use proportional racial balance as a standard could easily mean that the courts were indeed requiring "unreasonable" busing in many cities. The analysis had shown that there is a socially efficient domain of improvement in desegregation for modest increases or redirection in current busing provided that the busing were efficiently planned. Much of what was correctly perceived as "excessive busing" was the result of applying traditional bus routing techniques to a new problem which required a comprehensive method in order to avoid moving many children with little real reduction in racial isolation.

And the attempt to reach proportional racial balance in every school can lead to thousands of children being bused with little overall added gain in physical desegregation (see table 7; the increase in busing in moving from the fifth case to maximum desegregation).

HEW used these results to fight a two front policy battle. Against those who wanted to stop all busing for integration, the prospect of a reasonable middle ground was opened up by the analytic results which suggested that more efficient busing and sensible time limits on transportation could replace some of the burdens that had been imposed to date. At the same time, the analysis provided the first clear explanation for the widely perceived unreasonableness of the court ordered urban busing. The combination of small town bus dispatching methods and proportional racial balance standards--explicit or implicit-in fact meant that far more children than necessary were being ordered to move from their neighborhood schools. In fact, the prospects of reconciling the social benefits of attending neighborhood schools and desegregation were opened up by the

analysis results. In many of the cities examined, it became clear that if the school district tried, it could often substantially increase integration by making an effort which involved no increase in busing combined with a shift in school attendance zones. In the illustration (table 7) this is case #2 which improved the desegregation index from 44% to 68% using this neighborhood reassignment alone.

Because of the political sensitivity of the busing issue, this policy debate involved only a few people together with the Secretary in HEW and a few officials in the Executive Office of the President and the Department of Justice. The methods, findings and policy implications of the analysis were presented to these individuals. By the time the President had returned from his China voyage in February 1972 the tide had shifted away from an anti-busing constitutional amendment and the issue now became one of searching for some reasonable busing standards. That task was given to Attorney General John Mitchell and the Department of Justice. Their tendency was to propose such severe restrictions that urban desegregation would have been virtually impossible. In a sense, the federal policy debate about busing limits faced the same problem as the federal courts--the complexity of the facts made it impossible for normal intuition and common sense to predict the actual effects of limits that might on their face appear to be sensible. The exception to this of course, was the notion of a limit on the busing time-here there was agreement that roughly 30 minutes each way for elementary children and 45 minutes for high school students was a sensible maximum.

HEW was able to formulate a series of normative standards for busing and then simulate their effects in real cities on the twin objectives of policy--accomplishing desegregation with minimum necessary busing. This analytic capability resulted in effective HEW opposition to a number

of the most restrictive busing limits and provided several apparently reasonable alternatives for choice by the President.(13) Only this analytic capability permitted HEW to recommend "limits which prevent excessive busing but permit reasonable desegregation." (14)

The end result in terms of the decisions made by the President was decidedly mixed from my perspective. On the positive side, this large study and its use in the policy debates had clearly contributed to a much more constructive approach than had originally been planned. The evidence that progress in urban desegregation was possible without large scale increases in busing opened the way for the fundamentally important decisions to provide additional funds and incentives for reduction of racial isolation. This was much better than an entirely negative anti-busing action of one type or another. Further, the President decided upon busing limits as part of the Equal Educational Opportunity Act which were more compatible with desegregation than those originally proposed by others in the administration.(15)

At the same time, however, the President submitted a second bill to the Congress entitled The Student Transportation Moratorium Act of 1972. According to this bill, a moratorium on any new busing for desegregation was to last until July 1973 "in order to provide Congress time to fashion..a standard" for busing. It was probable that the capability presented by the HEW study to simulate the effects of different busing limits served to educate all participants about the complexity of the issue. This might have explained part of the decision to ask Congress to establish specific standards. Probably, though, politics played the key role--the submission of the Transportation Moratorium permitted the administration to intervene on the side of strong currents in public opinion and shifted responsibility for further action to the Congress.

Besides this three month policy debate, several other uses were made of this analysis work. HEW faced a number of imminent desegregation enforcement decisions involving various large districts. An unexpected benefit of the analysis was its usefulness as an indicator of whether there was illegal segregation. Whenever the desegregation index for the neighborhood school assignment (case 1, table 7) was substantially higher than that actually observed in a district where the amount of busing was greater, it was a clear sign that some of the busing was segregating children on purpose. At the same time, the "efficient busing" approach could permit HEW to reach voluntary agreements with districts based on the attractive deal that they would not have to undertake any additional busing provided they used existing busing to maximize desegregation. The analytic results opened this novel opportunity up and the Secretary and Director of the Office for Civil Rights decided to try this approach with a number of large districts.

But there were two areas of practical implementation where little use was made of this work. One of the original reasons for undertaking this effort was to provide HEW with an independent means of assessing whether large school districts with large minority enrollments had in fact approached the "maximum possible" desegregation effort required before they could be eligible for compensatory education funding under terms of the 1970 ESAA bill. Yet by the time HEW moved into its final decisions on implementation of ESAA in late 1972 there had already been two full years of "temporary" program experience. During that time, an allocation and priority ranking system had been devised which tried to focus the grants on school districts making the greatest efforts toward integration. This had led to the construction of several measures of effort based on the plans submitted to HEW and the program managers wanted to continue using this approach. The Office of Education did not fully understand or trust the 44 city analysis and did not want to rely on anything that complex in its assessment of

whether school districts were proposing "enough" desegregation to justify their receiving various kinds of funds. And, the analytic staff was already in the midst of enough other conflicts rising out of attempts to make use of evaluation results and focus resources, that this issue was easily won by the program staff.

The other major purpose of the busing study was to provide urban school districts with a new method for approaching this question. Toward that end, the initial contract had specified that a simple and explicit handbook be available so that school districts could apply the same methods to their schools. During testimony before the Congress in the spring of 1972 Secretary Richardson made explicit mention of the availability of this material and urged school districts to consider making use of the new approach in their desegregation planning. A large number of news stories about the report circulated during the time when it was being briefed to the HEW staffs after the two Presidential intentions had been sent to Congress. This exerpt from a Joseph Kraft column conveys the flavor:

> Behind the scenes in Washington there is now circulating an explosive document--entitled "School desegregation with minimum busing." Its findings...summarize what is probably the only systematic attempt ever made to study the ratio between busing of students and desegregation...Measured against the analysis..the new Nixon anti-busing package is not nearly as severe as Mr. Nixon has made it seem.(16)

In spite of this notoriety, and the endorsement of the new busing methodology by HEW officials, school districts displayed no interest in using the approach as part of a comprehensive desegregation planning process. If the ESAA program had announced its intended use of this as a measuring standard for grant proposals, there might have been more local interest. But my guess is that the traditional methods were entrenched and there was little interest in acquiring the skills needed to apply what appeared to be a complex new approach--even though it

might have involved considerable efficiency and financial savings. There was also a great deal of resistance to the notion of "computer simulation" or any use of the computer in drawing up alternate desegregation approaches. Finally, the basic premise of "desegregation with minimum busing" may have been inconvenient for those who constantly used and even exaggerated the burdens of busing as an argument for avoiding any real moves toward integration.

Relatively little use of any kind was made of the two other busing studies. The separate small contract for analysis of Los Angeles had been undertaken to test an alternative analytic methodology in case the more ambitious 44 city approach failed. As it turned out, the second method was far more cumbersome and much less flexible. However, the Los Angeles study was used along with the 44 city report in the policy debates to indicate that the results in both, separate, studies tended in the same directions. It served to verify the essential findings of the larger analysis and reinforced the case that metropolitan desegregation could be a positive alternative to urban isolation. But, that point of view had no impact on the final policy decisions in 1972.

The third small contract represented an effort to find a new and better method of urban desegregation planning without the use of computers and mathematical optimization techniques. This was intended to be helpful for medium sized districts-those with enrollments between 25,000 and 60,000. The results were inconclusive and there was little interest on the part of these districts in the hand methodology which was developed. Essentially, these two additional contracts represented small contingency investments in the event the major study was unsuccessful or behind schedule.

III. Achievement effects of compensatory education

In effect, the compensatory education issue became a central focus for the political, policy and analytic debates about all education policy during these three years. At the start of the 1970 decision making process, a very unusual coalition in the White House made the strong assertion that "compensatory education does not work." In his March 24, 1970 speech, the President had proposed a "quality education" portion of the desegregation assistance program and this coalition wanted to make certain that there would be no substitution of compensatory education for desegregation. The "southern strategy" people were very anxious to be certain that as much as possible of the $1.5 billion would flow to the southern school districts where nearly all of the federally required desegregation was occurring. They understood that most of the "racially isolated" minority children where compensatory education funds might be spent were outside their region. (17) At the same time, Dr. Patrick Moynihan and members of his staff were concerned that the President or school administrators might seize upon compensatory education as a way out of desegregation. They wanted to make sure tha there would not be analytic rationale for believing that compensatory education could do anything much for the education of minority children as long as they remained isolated.

It was in this context that members of the Secretary's planning staff undertook an examination of the available evidence. Fortunately, Dr. Alan Ginsburg, had been heavily involved in an earlier, separate review of compensatory education which had been undertaken by HEW as a result of the negative findings of the national evaluations which had been conducted up to that time. Those negative findings had surprised many observers, demoralized many in the Congress and HEW and were one of the major reasons for the rapidly spreading dogma that "compensatory education does not work". In an effort to get a

broader perspective on the issue, Dr. Ginsburg had assembled all the annual evaluations of compensatory education produced by the states and any other evaluations that might have been done. All of this evidence, a veritable mountain of reports and studies, was then carefully examined by three individuals who worked night and day to complete the project in time to be helpful in the policy debate. (18)

This synthesis of the available evidence was done with the same effort at comprehensiveness and objectivity that had been applied to the assessment of the educational effects of desegregation. Only evaluations which used pre-test and post-test achievement data were included and that narrowed the number of state evaluations from fifty to seven. A number of other large scale and smaller evaluations were also available and the positive and negative results, together with estimates of the reliability of the data, were summarized in a series of charts. That still left one important step---the use of judgement to draw conclusions from the mosaic of evidence.

There were three sources for the then prevailing view in the evaluation community that compensatory education had failed and could not help disadvantaged children. First, "success" had usually been defined as bringing a slow-learning child fully up to normal levels of performance. We decided that there were three categories of success: low gains in achievement when a student performed better that the level usually attained by disadvantaged children-above 7 months gain in reading during 10 months of school); medium achievement gains when the improvement was between 10 and 14 months improvement in reading; and, high gain-for anything above 14 months of improvement. In other words, any improvement in basic skills better than would usually have occurred without extra help was defined as a positive result by the HEW analysis.

The disappointing results of the national evaluation of the 1968-69 federal compensatory education program also looked different when a closer look was taken. Using HEW data it was possible to determine that less than 40% of the federal

compensatory funds had been spent for help with basic skills. Further, the funds were spread so thinly that the average participant received no more than a few dollars in educational services each month--not enough to make much difference in reading. This raised the question of whether there really had been much of a compensatory program in place in most schools evaluated. And, although the national evaluation had included nearly 100,000 students, pre-test and post-test achievement data had been collected for only a small proportion of the participants. These facts greatly reduced the persuasive impact of the achievement results. Perhaps one might say that a non-program deserved a non-evaluation!

A third major component of the negative consensus was the inference drawn from the 1966 report on Equal educational Opportunity. One of its important findings had been that variation in school resources was not a significant independent factor in explaining differences in the achievement of children. This finding had been coupled with the one mentioned earlier--that the social and economic background of children was an important factor in explaining school outcomes and specifically that children in schools more than 50% non-poor tended to achieve better. Moynihan had drawn heavily on these findings in his formulation of the Family Assistance Plan of 1969 and they made him feel that additional school efforts could accomplish little in majority poor schools.

While accepting the findings of the Coleman report the 1970 HEW analysis rejected the inference that these meant that no matter what is done in school there will be no achievement gains for slow learning children. The 1966 report had analyzed the existing varieties of school conditions at that time before any federal compensatory program existed. Therefore it had nothing definitive to say about situations in which substantial additional efforts above normal school expenditures, were provided for basic skills.

On the positve side a number of the state and local evaluation reports showed an impressive range of positive results. These evaluations often involved many thousands of pupils. Further, the state reports seemed to suggest a pattern of relationship between higher per pupil expenditures on basic skills and better results (see table 8 below).

Table 8 PROGRAM COSTS AND ACHIEVEMENT RESULTS - 1970

| Program | Achievement Results | Per Pupil Costs | % of all Pupils |
|---|---|---|---|
| California Title I (94,000 sample) | High achievement gain | $283 | 14 |
| | Low achievement gain | 160 | 26 |
| Colorado | High achievement gain | 207 | |
| | Medium achievement gain | 170 | |
| State Title I report | Low achievement gain | 122 | |
| Connecticut | High achievement gains | 305 | |
| State title I report | Medium achievement gain | 154 | |
| Analysis of successful compensatory education projects in California | High achievement gains | $237 (average) | |

By the third week in April 1970 the staff synthesis had been completed. A memorandum summarizing the evidence and attaching overviews of the positive and negative conclusions reached by all studies on compensatory edcuation achievement and costs came to this conclusion:

> "compensatory education has accomplished less than expected but more than is generally believed...More importantly, where funds were spent approximating $300 per pupil there is a consistent pattern of improvement". (19)

Taken together with the demographic analysis which showed that in many cities, even total desegregation would leave many minority children in schools with where most children were from low income homes, this suggested that HEW make a strong case for a substantial compensatory education effort as part of the program. Secretary Finch accepted this recommendation and requested that the analysis be presented at a meeting of the Cabinet committee on April 24, 1970. Naturally, the HEW analysts were delighted with the opportunity to do this and only after arriving in the White House did they discover that Patrick Moynihan had also prepared a memorandum taking the opposite point of view. In the end,and not necessarily because of the persuasiveness of the HEW analysis, the HEW position was adopted by the President and included in the ESAA bill.

As with the desegregation question, this experience with the patchwork of existing evaluations resulted in immediate efforts by the planning staffs to obtain more and better information about the effectiveness of compensatory education. Three steps were decided upon:

1. A large scale longitudinal study of the results of compensatory reading programs was designed and begun under external contract; results would not be available until 1974.

2. To provide more immediate results, a small contract was awarded for the location; evaluation, reanalysis and synthesis of all current evaluations of Title I or compensatory education programs that used pre-test and post-test achievement data.
3. Immediate efforts were made to improve the policy utility of the annual Office of Education evaluation of Title I programs. Despite the expenditure of nearly $2.5 million each year, this evaluation had failed to provide significant insights concerning achievement because school districts regularly reported usable achievement scores for only 7% to 8% of the children in the national sample.

Throughout 1970 and 1971 the evaluation results were used by HEW to encourage greater focus of the compensatory program on the basic skills (reading and mathematics) and the greater concentration of resources on fewer participants. This, together with new legislation designed to prevent local school districts from using compensatory funds to substitute for rather than add to their own resources, constituted the major HEW efforts to make the federal compensatory education program work better.

As analytic work began on the desegregation issue in early 1972, the compensatory edcuation question came up again in that context. In early February the planning and evaluation staff in the Office of Education was given the assignment of reviewing the evidence and its policy implications-. This work was to be completed within a very tight time schedule.

As the busing battle was resolved, I began to work with members of the Secretary's planning staff to think about the educational component of a new approach. In thinking about this, it seemed that a way had to be found to improve the education of poor and minority children who were isolated in predominantly poor schools in places where desegregation simply could not reach many of the children because of demography and existing

school district boundaries. The educational problems were greatest in urban school districts with large proportions of low income children and most of these districts also had enrollments where minority children outnumbered others. And, according to HEW evaluation data,these districts also received less than their fair proportion of existing compensatory education funds. Those that were received, were spent in the usual pattern--a little bit on many activities for many children with little focus on basic skills. Political realism also suggested that any new solution would have to rely on incentives and a good deal of local initiative.

It seemed that one way of achieving a number of federal objectives at the same time was to combine aspects of the 1970 desegregation assistance program with the purposes and resources of the federal compensatory education program to give school districts an incentive both to desegregate and to provide concentrated help for poor children. The new approach suggested to the Secretary included the following features:

- to be eligible a district had to be making progress with required or voluntary desegregation efforts and have one or more schools where children were racially or economically isolated;
- all poor children in schools with more than 25% low income enrollments would be eligible for a concentrated level of compensatory education assistance ($200-300 each) to be spent on basic skills:
- for each poor or minority child transferred from a majority poor or racially isolated school to a school with more than 60% white or middle income enrollment, the district would receive a transfer "bonus payment" for improvement in basic skills while retaining the compensatory education funds already being received for that pupil.

This design intended to bring together the two major strands of evidence from the education evaluations: poor children have the greatest learning difficulties in school and they are helped by being in schools where most of the children are from middle income homes; second, that with a focus on reading and mathematics together with a reasonable amount of extra resources, compensatory programs can help low achieving children even in economically or racially isolated schools. The proposal also affirmed the twin objectives of equal opportunity policy--- to provide federal encouragement both for desegregation and improved education for the least well off children. HEW estimates showed that four to five million low income children or, nearly half of those eligible for the current federal compensatory program, were enrolled in majority poor schools. By combining the desegregation requirement with explicit incentives for school districts to transfer some of these low income children into better learning environments where all reserach showed they would have a better opportunity, the proposal intended to shift the emphasis from racial balance to both racial and socio-economic integration.

On March 1, 1972 this was discussed at length with HEW Secretary Richardson. He found the concept of great interest as did his director of education planning and the Commissioner of Education (at that time I served as Assistant Director for the Office for Civil Rights). There remained the task of costing out the program ideas, thinking about different types of linkages with the pending ESAA legislation and the existing federal compensatory program, and, coming to some decisions about the current state of the evidence on compensatory education. (20)

Unfortunately the sythesis of evidence which was to have been available from the Office of Education had not been completed. The pace of decision was now gathering momentum and

I was asked to pull together the best available evidence and come to some judgement about the issue. It was at this point that a sharp dispute broke out between the analytic staff in the Office of the Secretary and the Office of Program Planning and Evaluation (OPPE) in the Office of Education. There were two reasons and both derived from evaluation studies. In 1970 the national Title I evaluation which was available to HEW covered the 1968-69 school year. In February 1972, preliminary results were starting to come in for the national evaluation of compensatory programs for the 1969-70 school year. These preliminary findings once again were negative and since the evaluation had cost two and a half million dollars--these were being given considerable weight.

The position I took was the same as with the previous national evaluation. The results did show that the federal program in compensatory education was not working to raise achievement. But once again, program funds, had been spread thinly and not invested primarily in reading and math, good test data was available for fewer than 9% of the pupils in the sample. Although the program as administered in 1969-70 might have failed again in educational terms, that did not mean that compensatory education as a process could never work. That elementary distinction seemed to escape many of the analysts debating the evidence.

A second evaluation which was also perceived very differently in the two analytic wings of HEW was the sythesis and analysis of state evaluations of compensatory education. The Office of Education quoted the conclusion of the report that "judging from the information given by those states, there appears to be little or no statistical evidence of a positive relationship between per pupil expenditures and project effectiveness" (21) That report had been eagerly awaited because the original intention was to have an independent team evaluate the methodology and data used by those states which actually

produced results based on matched pre-test and post-test data. The HEW group in 1970 had to take the state evaluation reports at face value and therefore hoped to get a more in-depth assessment of their reliability.

But that is not what the contractor did. For reasons never made clear, the overwhelming portion of the study effort was invested in a management review of the federal compensatory program. Thus, the OPPE sythesis reported that after looking at 91 state evaluation reports, only seven contained hard data on achievement and costs--- precisely the point made by the 1970 HEW analysis. In the end, it turned out that the outside education evaluation specialists had done the same thing that the HEW analysts had done two years earlier. They had not done any further digging into the available evaluations and the data was for the school years 1967-69.

As a result the question then came down to one of differing opinions about the meaning of the evaluation data from the states which did have test data. The contractor and the Office of Education choose not to be convinced by those results, while in 1972 as in 1970 the HEW analysts saw a pattern that seemed to be significant.

As matters turned out, I and several associates went to work to assemble the collage of information from all the sources available. This included the major national evaluations, state evaluations, reports done by cities and a large number of individual studies of compensatory program. The requirements for inclusion in this overview were that the study had used pre-test and post-test achievement data and had been conducted by a responsible institution.

Several of the state reports were very persuasive because they were carefully written, well documented and seemed to have worked out a thorough methodology for analyzing the achieve-

ment data from tens of thousands of children. Two examples were California and Colorado. For that reason we arranged to undertake several verification analyses of the California data and hired independent educational measurement specialists from several universities. They visited the state education agencies and were given permission to dig into the data and methods in order to determine whether these were valid and reliable. The verdicts were positive and by March 10, 1972 the synthesis was ready for presentation to the Secretary. (22) The conclusions were quite similar to those reached in 1970 but this time there was more documentation. In essence the conclusions were entirely probablistic and might best be summarized by quoting the President in his extensive discussion of this issue in the March 17, 1972 message to the Congress proposing the Equal Educational Opportunity Act: (23)

> For some years now, there has been a running debate about the effectiveness of added spending for programs of compensatory or remedial education. Some have maintained there is virtually no correlation between dollar input and learning output; others have maintained there is a direct correlation; experience has been mixed.
>
> What does now seem clear is that while many Title I experiments have failed, many others have succeeded substantially and even dramatically; and what also is clear is that without the extra efforts such extra funding would make possible, there is little change of breaking the cycle of deprivation.
>
> A case can be made that Title I has fallen short of expectations, and that in some respects it has failed. In many cases, pupils in the programs funded by it have shown no improvement whatever, and funds have frequently been misused or squandered foolishly....
>
> The most prevalent failing has been the spending of Title I funds as general revenue. Out of 40 States audited between 1966 and 1970, 14 were found to have spent Title I funds as general revenue....
>
> On the positive side, Title I has effected some important changes of benefit to disadvantaged children....

In designing compensatory programs, it is difficult to know exactly what will work. The circumstances of one locality may differ dramatically from those of other localities. What helps one group of children may not be of particular benefit to others. In these experimental years, local educational agencies and the schools have had to start from scratch, and to learn for themselves how to educate those who in the past had too often simply been left to fall further behind....

While there is a great deal yet to be learned about the design of successful compensatory programs, the experience so far does point in one crucial direction: to the importance of providing sufficiently concentrated funding to establish the educational equivalent of a "critical mass," or threshold level. Where funds have been spread too thinly, they have been wasted or dissipated with little to show for their expenditure. Where they have been concentrated, the results have been frequently encouraging and sometimes dramatic.

In a sample of some 10,000 disadvantaged pupils in California, 82 percent of those in projects spending less than $150 extra per pupil showed little or no achievement gain. Of those students in projects spending over $250 extra per pupil, 94 percent gained more than one year per year of exposure; 58 percent gained between 1.4 and 1.9 years per year of exposure. Throughout the country States as widely separated as Connecticut and Florida have recognized a correlation between a "critical mass" expenditure and marked effectiveness.

After the President sent the two legislative proposals to the Congress on March 17, 1972, it was necessary to prepare briefing materials and testimony for the hearings which would be held in the House and Senate. In order to clarify the existing evaluation evidence on compensatory education I was asked to prepare a more extensive presentation for distribution to the public. Following the end of the congressional hearings in late March I and two colleagues began the further collection of evidence and analysis.

Independent education measurement specialists were again hired to visit additional states and major cities which had produced evaluations using achievement test and program cost data. Again, the purpose was to examine the validity of the methods and raw information used.

Due to the public debate and controversy surrounding the proposals of the President, most of which centered on the busing rather than the education question, there was a need to finish this synthesis as soon as possible. Within sixteen days a first draft report was ready for review by the Secretary and others in HEW. It had been almost two years to the day since the encounter with Dr. Moynihan on this issue in the cabinet committee meeting and I described his opposition to the HEW conclusions at that time. Secretary Richardson decided to ask Professor Moynihan to review the draft report and a copy was brought to him at Harvard University. He informed the Secretary that the report was comprehensive, treated the evidence fairly and that he could support the overall conclusions reached. The Office of Education planning and evaluation staff, however, continued to have some reservations about the report and its conclusions.

Nevertheless, on April 20, 1972 the Secretary held a press conference to explain the report and make it available to the public. At that time he repeated that the 207 page HEW "white paper", The effectiveness of compensatory education, summarized the best available evidence and reached the conclusion that with the right kind of local effort compensatory education could be made to work. Further, he repeated his congressional testimony and quoted from the report in saying that "compensatory education and desegregation are complementary" approaches to better education for all children. (24) In response to a reporter's question about whether there was sufficient evidence to reach a responsible policy conclusion, Secretary Richardson replied: "with this much evidence before us, we would be irresponsible not to come to a conclusion about what it means". In essence the intense timetable and pressures of the policy process from January through March 1972 forced HEW to confront the evidence, make specific proposals to the President which he accepted, and then share the evaluation data with the public in

in an easily understandable form. The policy conclusions were consistent with the trend of HEW efforts to improve the impact of the federal compensatory program since 1969.

EEOA

The new program proposed by the President would have committed an additional billion dollars per year on a continuing basis. This in combination with current Title I expenditures of roughly $480 million in the school districts likely to be eligible and expected annual funding for ESAA at $500 million for three years, combined to form the single largest HEW federal education program initiative. The HEW recommendations for the substance of the program were accepted by the President and he described the purposes in these words:

> The program which I propose aims to assure schools with substantial concentrations of poor children of receiving an average $300 compensatory education grant for each child.
>
> In order to encourage voluntary transfers, under circumstances where they would reduce both racial isolation and low-income concentration, any school accepting such transfers would receive the extra $300 allotted for the transferring student plus a bonus payment depending on the proportion of poor children in that school
>
> One key to the success of this new approach would be the "critical mass" achieved by both increasing and concentrating the funds mad available; another would be vigorous administrative follow-through to ensure that the funds are used in the intended schools and for the intended purposes. (25)

No action was taken by the Congress, although EEOA was reintroduced and discussed until 1974. Since then both the racial and the economic isolation of low income students in most urban school districts has increased and the Supreme Court has ruled that compulsory metropolitan area integration goes beyond the Constitution. Six years later the question of effective approaches to help millions of these young people learn the basic and necessary skills remains before the nation.

D. Factors affecting the use of evaluations

The use of the evaluations which has been described in this review was far from the typical pattern. But for those evaluations studies which are used for real decisions inside the federal government, there probably is no "typical pattern". The annual budget and program review cycles are supposed to be the occasions when regular and systematic use of evaluation reports is made. But all experienced observers of the federal government know that those are rarely the times when basic decisions are actually made. There are generally too many constraints imposed by the combination of constituency group influence, the weight of things as they are and the Congress. Therefore, it may be that the most salient occasions when social science evaluations can have some impact in the world of action may resemble these three hectic years----events or the purposes of the political leadership combine to determine the need for a new approach to perceived problems. Three different types of factors combined to affect the extent to which these evaluations were used to reach decisions; these were institutional, personal and situational.

Institutional

Political leaders tend to be very practical and instrumental individuals. A logical question for them to ask when examining a course of action in any area of public policy is---what do we know about what works to accomplish the purposes we have in mind? In 1970, the new craft of program evaluation had already been in existence for several years along with small staffs which had the task of interpreting the studies. This established the habit and legitimacy of turning to the evaluation reports for an assessment of progress to date. That simple fact, along with the extensive use of evaluations by the White House and HEW staffs in the first year of the Nixon administration established the pattern of incorporating this strand of insight into much of

the decision making. And, by the same process, once the available evaluations had been used once in 1970, it was nearly axiomatic that these would be given at least a hearing, even in the more highly volatile atmosphere of 1972.

In 1970 there was a rather well structured decision making process which linked the White House and HEW together in a predictable way. This permitted different points of view---based on quite varied political and factual perceptions--- to compete and use evaluation derived information in doing so. The situation was more amorphous in 1972 but the number of participants was so limited that the analytic point of view remained visible. By contrast much of the real decision making on implementation took place in a way designed to give greatest weight to bureaucratic maneuvering and insider access while excluding information coming from outside the program itself. There was no defined calendar and process for making many of the implementation decisions in contrast with the tight time table established for both episodes of policy decision.

Of great importance was the role that had evolved for the analytic staff of the Secretary in HEW. This small group of individuals was at the intersection of decision-making and evaluation with full participation in each of those very separate spheres. This made it possible for the evaluation agenda in subsequent years to reflect the new questions and priorities that emerged from the 1970 policy making experience. And, at the same time, when the evaluations became available, this staff could immediately relate them to the policy context. The Commissioner of Education did not tend to use the planning and evaluation staff to the same extent--partly because more of his concerns were with management and budget issues. That resulted in a far greater "social distance" between evaluation research and federal decisions in fact and in the perception of that planning staff. To some degree, this may explain much of the

reluctance of that staff to really make efforts to apply the results of evaluation to actual policy decisions. Many seemed to view this activity as "dirty" because it was not "scientific".

The combination of tight time pressures and perceived importance of the decisions in both policy making rounds, worked to limit the number of participants. It also created a sense of shared burdens and interdependence among participants that undoubtedly increased the role of the analytic staff participating in the enterprise. Of course, these same institutional factors can work in precisely the opposite direction -to the exclusion of any and all evidence and analysis. Where the leadership had already come to understand the methods, contributions and limitations of the analysts, and had come to trust the judgement of the individuals, these factors would tend to heighten the use made of knowledge from evaluations.

The case studies demonstrated how little use was made of the evaluations in deciding upon program implementation. One major institutional reason is the widely understood distinction between aggregate results and the facts of particular situations. For example, although the ESAP I evaluation might have identified the three most effective desegregation assistance activities in the aggregate, it was understood that this might or might not apply in a particular school district applying for funds. Therefore, when the program manager insisted that those judgments remain open based on the particular situation--he had a strong case. This and basic notions of managerial accountability make it difficult to impose evaluation derived implementation rules on reluctantmanagers. It is also difficult to transpose evaluation findings into procedures for day to day action--especially when these derive from complex studies which are open to challenge. The unwillingness to use the 44 city results in the ESAA program was an example of this.

Personal

We all know that individuals give life and content to any institutional setting. Although the leadership of large institutions like HEW faces definite limits on what is possible in many domains of action, their own personalities and commitment can distinctly shape some events. This is probably most true in all areas of policy decision and undoubtedly explains why activist cabinet and policy officials tend to invest most of their time and energy on what are perceived as unique policy opportunities rather than management and bureaucratic performance questions.

Both in 1970 and in 1972 HEW was led by dedicated, problem-solving individuals who were eager to learn about and use the type of evidence that evaluations could provide. In 1970 this group consisted of Secretary Robert Finch, Undersecretary John Veneman and Lewis H. Butler, Assistant Secretary for Planning and Evaluation. They provided the sense of direction and the political judgment which permitted the analytic evidence to become an integral element of all decisions made by the President. In 1972 the equivalent group consisted of Secretary Elliot Richardson, General Consul Wilmot Hastings and Marshall Moriarity, Executive Assistant to the Secretary. Unless these individuals had been willing to take policy initiatives, they would not have asked their staff for alternatives or proposals. Unless, they had been willing to go against the prevailing popular and White House opinions, they would have had no need for any of the products of the evaluation process. And without their capacity to actively understand the often complex data and creatively use it for the decisions to be taken, there would have been little policy utility in the evaluation research. These individuals also understood and accepted the professional responsibility of the policy analyst to recommend with total objectivity as his best judgement indicated.

If the results or proposals were not in keeping with their political preferences or instincts they would decide as they thought best, but they would never try to compromise the analyst or the evaluation process by bending or suppressing evidence.

The good personal cooperation at the staff level during these years in HEW was also another major factor. In this area, individuals worked together with a focus on issues rather than turf and competition. That made it possible for the top leadership to concentrate on the substance rather than spend half the time sorting out the hidden power plays. The good cooperation given by Dr. John Evans, leader of the planning staff in the Office of Education, is one of many examples. Unless he had agreed to fund a number of evaluation studies that had derived from the ESAA planning process, hardly any of the new information would have been available in 1972. Not only did he provide the funds and management for these evaluations, but Dr. Evans also had a strong interest in seeing that these were used.

A last personal factor is that a number of the individuals with responsibility for analysis and evaluations were willing to use their admittedly fallible judgement to draw conclusions from ambiguous evidence. And, these staff members combined a scholar's understanding of the social science methods with an instint for how the conclusions might be translated into specific suggestions for federal action.

Situational

At first there is a temptation to say that the reemergence of the issues, at the Presidential level where they had to get attention, within a two year period is a historical accident. Indeed, that situation had a lot to do with the use of evaluations. However, most large issues in public policy and especially in education do reemerge again and again---until society solves the problems or loses hope. Therefore the pattern may be more the rule than the exception.

More unique was the continuity of some of the participants over a three year period. The fact that several of the policy analysts were still working on these issues (see table 4- Evans, Ginsburg, Menges, Timpane) provided an important connection from the issues to the evaluations and from the evaluations to policy recommendations. Unique, in a negative sense, was the fact that the two years of temporary desegregation assistance programs, essentially worked to define the management approach in such detail that there was little interest in anything the evaluations might have to offer. Although, the existence of the evaluations and the fact that several analysts were highly critical of the first year of ESAP had served to improve the management arrangements HEW used.

RECOMMENDATIONS

Social science evaluations of public programs have been required by the Congress and supported with tax dollars in order to improve federal policy and program management. That simple purpose brings us to the question of how to improve the linkage between this type of knowlege and federal actions.

The cultures of social science research and public action have very different requirements and incentives. Public action intrinsically calls for simplicity, clarity, broad choices and a good measure of inspirational certainty. Research involves specific methods, focus, precision and cautious hedging of the results. Almost by definition, the more carefully controlled and replicable a social science research project, the smaller will be the domain of action for which the results will be directly and obviously applicable. Clearly, these worlds attract different kinds of people, with distinctive skills and varied perceptions of the contribution good social science evaluations might make. Communication and interaction require individuals and organizations that can command the respect and trust of both worlds and bring them together.

Too often in the normal course of events evaluations are either ignored or they are used as in an adversary proceeding to cancel each other out much like expert witnesses in a trial. Because the evaluations are often large, difficult and complex undertakings in themselves, it becomes too easy for the evaluation staffs to become so absorbed in getting their job done that they have little time left to think abouts its significance. Further, this functional specialization often means that members of the evaluation community spend far more time assessing and critizizing the methods, or techniques used in evaluations than in exploring their implications. This in turn forces the evaluation project officers inside the government to spend a great deal of time explaining or defending the evaluation

rather than making use of the results. And, since the natural tendency of organizations is to wall off planning and evaluation staffs from decisions, this process of encouraging the government evaluators to talk mostly with social research evaluators can become a tempting way to diminish the interference that might come from their findings.

In order to make use of evaluations three fundamental steps need to be taken. First, a decision must be made concerning the methodological validity of each specific evaluation report. This is basically a technical matter. Second, the valid evaluations which exist about a subject or program must be assembled in a cumulative way, their conclusions must be summarized and finally, a net overall assessment must be made. This is a task of synthesis and judgement----it is a task too rarely attempted inside the government. Without this overall conclusion, it remains possible for everyone to quote their favorite evaluation i.e. the one that agrees with them or to dismiss any and all evaluations as not representative of the whole universe, defective in approach and so forth.

The essence of all scientific progress is the accumulation of evidence. Yet the leap from the display of the results of a number of individual evaluations to some view about the overall meaning is intrinsically a matter of individual judgement. There is no strictly "scientific" approach although the integrity of the individual permits a strictly objective approach. There is a technical dimension to the accumulation of evidence--that is formal procedures to assure comparability among the studies. (26) But even after that has been done, there still remains an act of interpretation. Only when the evaluation staffs inside the government demand this of themselves, on a regular basis, will there be a start in the more widespread use of this enormous resource.

But there is a third task---once the overall assessment has been reached, a further question remains to be answered--what do the conclusions mean for federal action? What will or should the government do differently because it now has this evaluation data? Beginning in 1972, Dr. John Evans in the Office of Education made a good effort to extract the "policy implications" from the larger evaluation reports as they arrived. These "policy implications memoranda" constitute a good start in this direction even though they applied to one study at a time.

These two difficult and essential links between evaluations and action involve the use of judgement. This is craft not science and a skill likely to be improved through use and over time. In the cases described in this report, the HEW analysts produced four syntheses of evaluation results and proposed actions as a result. There was an increasing facility with practice. Yet, it is evident that the process would have benefited by more time and above all more collegial dialogue about what the evidence said and meant.

HEW should annually pick two or three policy issues and one or two implementation issues and structure a two-step process to make use of the evaluation results. The planning and evaluation staff should decide on a set of technically valid evaluations relevant to the particular question- for example, do compensatory reading programs work? Then a summary of the main findings of those evaluations should be made and a small group should meet to discuss and debate the overall meaning of the evaluations. This group should include senior officials of the evaluation staff, the program staff, the legislative staff and a representative of the Commissioner or Secretary. The style should be that of a problem solving seminar and ideally there should be no immediate decisions pending. After a thorough discussion, each participant should summarize his overall interpretation of the conclusions reached by the evaluations. The final written interpretation, however, would be the sole responsibility of the evaluation staff.

A similar process should be followed for the next step---the specification of what the evaluation results mean for a c t i o n. Again, this is an interpretive task where the evaluation staff would gain by having a dialogue with senior education officials who bring other kinds of experience and insights to the question. Undoubtedly their views would reflect the practical context and the discussion would give the evaluation staff member an enhanced ability to write a policy implications memorandum that would have some force.

If this type of dialogue became a regular aspect of government life for two ro three significant issues a year, it would certainly sharpen the interpretive skills of the evaluation staff. It would also bring evaluations out of the research community and right into the active experience of other education officials. It would broaden their thinking, make them aware of the evaluations which had been done and lay the basis for far more informed use of evaluations when urgent policy program or budget decisions provided the opportunity. At those times, the shared experience of these two or three "evaluation action implications" meetings held each year would have prepared the basis to do the same things when the two interpretive steps would in fact result in tangible decisions.

I want to emphasize that this recommendation calls for entirely consultative participation. That is, the evaluation staff must retain the independent responsibility to interpret the net meaning of evaluations about an issue and the implications for action. This has to be the case in order to assure the continued independence of the evaluation staff--otherwise the normal pressures of the program point of view would overwhelm any possibility of the genuinely objective use of those studies.

Generally, the suggestions for increased participation have involved the wrong end of the evaluation process--the design and implementation phase. It is quite natural that individuals being evaluated will try to control the findings if they are participants in the evaluation itself. And, since doing an evaluation well requires a great many specialized skills, the participation of people who lack that training will generally reduce the quality of the result. My recommendation involves consultative participation at the other end of evaluation---the time when the difficult interpretive judgements need to be made and when the experience of policy leaders and program managers is directly relevant to action implications.

In schematic terms the first interpretive step would require the evaluation staff to prepare a summary for the discussion with their federal colleagues that might look like this:

Policy issue: has the federal compensatory program helped poor children improve their reading skills?

| EVALUATION | SUMMARY OF FINDINGS + | - | 0 | WEIGHTING FACTORS | NET CONCLUSION |
|---|---|---|---|---|---|
| A. | | | | eg. extent to which the program had been implemented; sample and test data characteristics etc. | |
| B. | | | | | |
| C. | | | | | |
| D. | | | | | |
| etc. | | | | | |
| Overall Findings | | | | | |
| 1 | | | | | |
| 2 | | | | | |
| 3 | | | | | |

Initially the presentation would leave the net conclusion about each evaluation as well as the overall judgement blank. After discussion of the evidence, the evaluation staff would be responsible for reaching a net conclusion about each evaluation and formulating the overall findings from the sum of the evaluation evidence.

The next intepretive step, where the views and experiences of program and legislative liaison officials can be especially useful to the evaluation staff might be illustrated by the following schematic:

| MAJOR FINDINGS OF EVALUATIONS | IMPLICATIONS FOR POLICY/PROGRAM DESIGN | IMPLICATIONS FOR PROGRAM MANAGEMENT |
|---|---|---|
| 1 | | |
| 2 | | |
| 3 | | |
| 4 | | |
| etc. | | |

The evaluation staff would initially provide only the information on the findings. After the dialogue, which would surface different reactions to what those findings might or might not imply for action in the real world, the evaluation staff would have the task of arriving at specific recommendations. This process is not intended to reach consensus but rather to provide a forum for communication and inform the judgement of those who must interpret the evaluations. It is a design for institutionalizing the actual intellectual process that must occur in any effort to make practical use of the evaluations. Conflict with other points of view remains inevitable once the evaluation staff makes a serious effort to use these recommendations to make changes in policy or management. The Secretary or Commissioner of Education will have to hear the competing arguments and make a decision.

Connecting the reality of program efforts and evaluations of their success

The basic concern of social science evaluation methodology has been focused on meeting the usual criteria for scientific experimentation. But as was the case with the two national

evaluations of compensatory education, a recurrent problem is that there is no program in place to be evaluated or that there are such different degrees and types of implementation that these really become very important as research parameters. "Failures" then often simply mean that the program being investigated has not really started yet;

Therefore, there is an urgent need for methodological invention in the design of evaluations which will permit conclusions to be drawn from sub-sets of the evidence based on practical data about what is really happening. This is especially needed in the case of new federal programs where the customary delays in congressional appropriation of funds, start-up time and the like, often mean that the first year is not going to run smoothly at all. Much of the understandable hostility toward evaluation on the part of public managers stems from the fact that they have observed an unfortunate paradox: public and congressional interest is usually greatest at the start of new programs, while performance is going to be lowest. Once an evaluation concludes "no effects observed" after the first year, this weakens the support and continuity needed to actually achieve some cumulative impact. This is often described as "tearing a newly planted tree up to look at the roots in order to make sure it is growing".

This syndrome, has lessened partly because of greater humility and learning about the evaluation process and partly because less attention has been given to evaluation results. Building in methods to disaggregate evaluation findings into components for different levels of program effort would help to increase the utility of evaluations. Good examples of this approach were the evaluation reports on compensatory education issued by the state of California and number of large cities.

Timing

Ways must be found to discipline the timetable of evaluation at both the start and at the finish. The ESAP I evaluation lacked achievement test data because the evaluation funds were not available until August 1970 and therefore a contractor could not begin work until well into that school year--too late for any pre-test data collection. With that lesson in mind, every effort was made to make certain that the ESAP II evaluation could begin early enough to avoid repeating the problem. It did not and collected only post-test data as did the even more elaborate third year (ESAP III) evaluation.

A large part of this problem is outside the executive branch and rests with traditional lateness in congressional action on the national budget. In that sense there is a poetic equivalence in the fact that this late funding not only damages the quality of federal programs, because schools and colleges never can plan with certainty, it also often fatally undermines the evaluations of those programs. But some part of the problem also resides in HEW---often with the slowness and complexity of the procurement process which serves as a massive bottleneck year after year. Though, to be fair, the evaluation designs tend to land in the contracts office in a large cluster, year after year.

But, in my observation, it is the rule rather than the exception that evaluation contractors fail to meet their schedules. Partly this is because they view the task as "research" and do not fully understand how the results might be useful. As demonstrated in this report, there were several occasions when the lateness or mis-direction of evaluation studies deprived the government of information that could have been very helpful for large social policy decisions. HEW must identify those

evaluations that might have exceptional importance for future decisions, make extra efforts to manage them closely and incorporate some financial incentives and sanctions to motivate better "on time" performance. Also, the more the members of the federal evaluation staff are actually involved in decision making, the more they can and will communicate the need for timely and predictable delivery of results to the contractors who, after all, want their professional efforts to have a tangible impact on public policy.

FOOTNOTES

(1) U.S. OMB, Special Analysis Budget of the U.S., 1975, p.108

(2) J. Wholey, Federal Evaluation Policy, Washington, D.C. The Urban Institute, 1970

(3) President Richard M. Nixon, School desegregation: a free and open society, policy statement delivered March 24, 1970

(4) Meyer Weinberg, The Education of the Minority Child, Chicago: 1970

(5) Nancy St. John, "Desegregation and minority group performance", Harvard Educational Review, February 1970

(6) Equal Educational Opportunity, Washington, D.C. Government Printing Office, 1966

(7) In the early 1960s more than 65% of all minority children were poor; by 1970 the situation had improved and 38% of all minority children were poor. The interchangeable use of the terms "Poor" and "minority" child was not generally appropriate, but felt to be relevant to the urban schools with heavy minority enrollments

(8) These proportions are based on analysis of Office for Civil Rights data for 1969 through 1971.

(9) Mainly, the late funding of the small program by the Congress in August 1970 prevented the evaluation planners from making the very detailed arrangements needed with school districts before any achievement data could be obtained

(10) Memorandum to the Director, Office for Civil Rights from C.C. Menges, Desegration: Educational and Social Effects, the Evidence, February 1, 1972

(11) Not scheduled for completion until June 1972. But if the study had been on schedule, preliminary results would have been available.

(12) Secretary Elliot Richardson, Testimony to the Committee on the Judiciary, U.S. House of Representatives, April 13, 1972, p. 6

(13) Memorandum to the Secretary of HEW from C.C. Menges, Current Desegregation Policy Issues, March 5, 1972. This memorandum included the proposed design for an entirely new approach to equal educational opportunity as well as an analysis of five types of limits then being proposed by the Department of Justice and a proposal for "reasonable limits"

FOOTNOTES - cont'd.

(14) Ibid, tab C

(15) R.M. Nixon, Message to the Congress of the United States. The White House, March 17, 1972. with attachments, Message on busing, fact sheet, 3pp

(16) The Washington Post, April 4, 1972

(17) The regional consequences of alternative program emphasis were made explicit in a briefing memorandum prepared for the Cabinet committee, Overview of desegregation assistance program alternatives, H.E.W. memorandum April 14, 1970

(18) These included Alan Ginsburg, Susan E. Smith and the author.

(19) A.Ginsburg, S.Smith, C.Menges, A summary of the educational achievement results of integration and compensatory education, April 24, 1970, H.E.W. memorandum

(20) A first draft of the proposal with cost estimates and funding mechanisms can be found in: C.Menges to the Secretary, Current desegregation policy issues, March 5, 1972, H.E.W. memorandum

(21) OPPE, Office of Education, H.E.W. Data on the relationship of educational expenditures, student learning environment and student achievement, March 16, 1972

(22) C.Menges to the Secretary, Compensatory education: when it does and does not work, H.E.W. memorandum, March 10, 1972

(23) President Richard Nixon, Message to the Congress (on equal educational opportunity) March 17, 1972 pp 12-13

(24) C.Menges et al. The effectiveness of compensatory education: summary and review of the evidence, Washington, D.C: Government Printing Office, 1972

(25) President Nixon, March 17, 1972 op cit p 14

(26) Light, R.J. and Smith, P.V. "Accumulating evidence: procedures for resolving contraditions among different research studies", Harvard Educational Review, Vol 41 no.4, November 1971

## SELECTED BIBLIOGRAPHY

A. *Evaluations in the context of decision-making*


Cohen, D., "Politics and Research: Evaluation of Social Action Programs in Education," Development of Education Research, pp. 213-238.

Howe, H. *Policy Studies: Blueprints for Progress--Or Social Security for Social Scientists.* Presented to a meeting of the Association for Institutional Research, Wash. D.C., May 6, 1974.

Lange, Robert R., "Evaluation as Management Methodology," in *A Strategy for Evaluation Design*, Caspar F. Paulson, Jr., ed. Oregon State System of Higher Education, Sec. VII (1970).

Myerson, M. and Banfield, E.C. *Politics, Planning and the Public Interest.* The Free Press, Glencoe, Ill., 1955.

Rossi, H. and Williams, W., eds. *Evaluation of Social Programs--Theory, Practice, and Politics.* Seminar Press, N.Y., 1972.

Steele, S. *Contemporary Approaches to Program Evaluation: Implications for Evaluative Programs for Disadvantaged Adults.* Eric Clearinghouse on Adult Education, Syracuse, N.Y.

Wholey, J.S., *et. al. Federal Evaluation Policy.* Urban Institute, Wash. D.C., 1970

Williams, W. *Social Policy Research and Analysis: The Experiment in the Federal Agencies*, American Elsevier, N.Y., 1971.


B. *Methodological issues in conducting social program evaluation*


Alkin, M.C. *Evaluating the Cost Effectiveness of Institutional Programs.* Center for Study of Evaluation, U. of Cal. at LA, Los Angeles, Cal., 1969.

American Institutes for Research. *Evaluative Research: Strategies and Methods.* Pittsburgh, Pa., 1970.

Campbell, D., "Reforms as Experiments," in *American Psychologist*, Vol. 24, No. 4 (April, 1969), pp. 409-429.

Caro, F., Ed. *Readings in Evaluation Research.* Russell Sage Foundation, N.Y., 1971.

Committee of the Social Science Research Council. *Social Experiments for Planning and Evaluating Social Interventions* (Draft), Aug. 1973.

Hyman, H. and Wright, C.R., "Evaluating Social Action Programs" in Lazarsfeld, P; Sewell, W. and Wilensky, H., eds. *The Uses of Sociology.* New York, Basic Books, 1967.

Suchman, E.A. *Evaluative Research: Principles and Practice in Public Service and Social Action Programs.* Russell Sage Foundation, N.Y., 1967

Trow, M. *Methodological Problems in the Evaluation of Innovations.* Center for the Study of Evaluation, U. of Cal. at LA, Los Angeles, Cal., 1969.

Zaltman, G.; Kotler, P.; Kautman, I. *Creating Social Change.* New York, Holt, Rinehard & Winston, Inc., 1972.

C. *Using evaluations to suggest future policy directions*

Etzioni, Amitai, "Shortcuts to Social Change?" *The Public Interest*, No. 12 (Summer, 1968), pp. 40-51.

Menges, C. *The Effectiveness of Compensatory Education: Summary and review of the Evidence*, United States Department of Health, Education, and Welfare, Wash. D.C., 1972.

Rivlin, A. *Systematic Thinking for Social Action*. Brookings Institution, Wash. D.C., 1971

Szanton, P. *Systems Problems in the City*. The New York City-Rand Institute, N.Y., 1972.

Weiss, R. and Rein, M. "The Evaluation of Broad-Aim Programs: A Cautionary Case and a Moral" in *Readings in Evaluation Research*. Russell Sage Foundation, N.Y., 1971.

Williams, W. "Developing an Evaluation Strategy for a Social Action Agency" in *Journal of Human Resources*, 1969, pp. 451-465.

Levitan and Taggert, The Promise of Greatness, 1976